PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 5/04, 5/10, 5/16, 5/22, 15/82, A01H 5/00	A1	(11) International Publication Number: WO 99/02655 (43) International Publication Date: 21 January 1999 (21.01.99)

(21) International Application Number: PCT/US98/14109

(22) International Filing Date: 7 July 1998 (07.07.98)

(30) Priority Data:
08/889,655 8 July 1997 (08.07.97) US

(71) Applicant (*for all designated States except US*): UNIVERSITY OF KENTUCKY RESEARCH FOUNDATION [US/US]; 207 Administration Building, Lexington, KY 40506-0032 (US).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): LUSSO, Marcos, Fernando, Godoy [BR/US]; 2020 Armstrong Mill Road #2004, Lexington, KY 40515 (US). CHAPPELL, Joseph [US/US]; 1808 Bimini Road, Lexington, KY 40502 (US).

(74) Agent: ELLINGER, Mark, S.; Fish & Richardson P.C., P.A., Suite 3300, 60th South 6th Street, Minneapolis, MN 55402 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: PROTEIN KINASES AND USES THEREOF

(57) Abstract

Nucleic acid molecules are disclosed that are induced upon pathogen invasion or elicitor treatment. Such molecules are functional in plants, plant tissue and in plant cells for inducible gene expression and altering the disease resistance phenotype of plants. Such molecules are, or are related to, sequences of calcium dependent protein kinase genes. Also disclosed are methods for obtaining transgenic plants containing such nucleic acid molecules and methods for using such molecules. Polypeptides encoded by such nucleic acids are also disclosed herein.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

PROTEIN KINASES AND USES THEREOF

Statement as to Federally Sponsored Research

The research reported herein was performed in part with funding from the National Science Foundation of the United States Government. The United States Government may have certain rights in this invention.

Field of the Invention

This invention relates to nucleic acids encoding calcium dependent protein kinases, polypeptides produced from such nucleic acids and transgenic plants expressing such nucleic acids.

Background of the Invention

In plants, disease resistance to fungal, bacterial, and viral pathogens is associated with a plant response termed the hypersensitivity response (HR). In the HR, the site in the plant where the potential phytopathogen invades undergoes localized cell death, and it is postulated that this localized plant cell death contains the invading microorganism or virus, thereby protecting the remainder of the plant. Other plant defense responses include the production of phytoalexins, the production of lytic enzymes capable of averting pathogen ingress and modifications to cell walls that strengthen it against physical and/or enzymatic attack.

The HR of plants can include phytoalexin production as part of the response to invading microorganisms. For example, tobacco *(Nicotiana tabacum)* produces sesquiterpenes in response to microbial invaders, e.g., *Pseudomonas lachrymans*.

A variety of compositions can serve as elicitors of plant phytoalexin synthesis. These include one or

more toxic ions, e.g., mercuric ions, other chemically defined compositions, metabolic inhibitors, cell wall glycans, certain glycoproteins, certain enzymes, fungal spores, chitosans, certain fatty acids, and certain oligosaccharides derived from plant cell walls. See, e.g., Sequeira, L. (1983) *Annu. Rev. Microbiol.* 37:51-79 and references cited therein. Cell wall fragments of certain *Phytophthora* species and cellulase from *Trichoderma viride* but not *Aspergillus japonicum* pectolyase can also elicit the HR. Attack by other plant pathogens or an avirulent related strain can also induce the HR.

Elicitins are proteins produced by plant pathogens and potential plant pathogens. Elicitins can induce the HR in plants. Generally, but not necessarily, localized cell death is the result of the elicitin-induced response in the infected (or challenged) plant tissue. These responses mediate full or partial resistance to destructive infection by the invading, potentially plant pathogenic microorganism. Amino acid and nucleotide coding sequences for an elicitin of *Phytophthora parasitica* have been published. Kamoun et al. (1993) *Mol. Plant-Microbe Interactions* 6:573-581.

Plant pathogenic viruses including, but not limited to, Tobacco Mosaic Virus (TMV), induce the HR in infected plants. Bacteria that infect plants also can induce HR and thereby disease resistance; representative bacteria eliciting HR include, e.g., *Xanthomonas spp.* and *Pseudomonas syringae*. Plant pathogenic fungi generally do not induce the HR response after attack on a host plant, e.g., *Phytophthora parasitica* and *Peronospora tabaci* on tobacco hosts, but can induce the HR after attack on a non-host plant.

The signal transduction mechanisms involved in expression of disease resistance are under investigation

and some of the genetic and biochemical features have been outlined. See, e.g., Staskawicz, B. et al., Science 268:661-667 (1995). However, many aspects of signal transduction pathways and the role of many specific components are not well understood.

There is a long felt need in the art for methods of protecting plants, particularly crop plants, from infection by plant pathogens. Especially important from the standpoint of economic and environmental concerns are biological or "natural" methods rather than those which depend on the application of chemicals to crop plants. There is also a need in the art for plant polynucleotide sequences for enhancing and/or improving disease resistance in plants.

Summary of the Invention

Nucleic acids of the present invention are based on novel calcium dependent protein kinase (CDPK) genes and their corresponding proteins. Induction of expression of these novel CDPK genes is surprisingly rapid, i.e., mRNA transcription of such genes can be observed as soon as 30 minutes after elicitor-mediated induction of plant defense responses. Thus, the novel genes disclosed herein are among those genes that are most rapidly induced in response to signals indicating an invading plant pathogen.

An isolated polynucleotide is disclosed herein, that comprises the nucleotide sequence of SEQ ID NO:1 and its complement, and an RNA analog of SEQ ID NO:1 or its complement. Such a polynucleotide can also be a nucleic acid fragment of the above that is at least 20 nucleotides in length and that hybridizes under stringent conditions to genomic DNA encoding the polypeptide of Figure 3. The polynucleotide can comprise, for example,

nucleotides 1 to 170, nucleotides 160 to 560, or nucleotides 550 to 920 of Figure 2.

A nucleic acid construct as disclosed herein comprises a polynucleotide of the invention. In such a construct, a polynucleotide of the invention can be operably linked to one or more elements that regulate transcription of the polynucleotide, for example, a regulatory element induced in response to a plant pathogen such as a fungus (e.g., *Phytophthora*), a bacterium (e.g., *Pseudomonas*), or a virus (e.g., Tobacco Mosaic Virus) as described herein. In other embodiments, such induction is mediated by an elicitor (e.g., by fungal or bacterial elicitors).

Further aspects of the present invention are transgenic plant cells, plant tissues, and plants that have been genetically engineered to contain and express a polynucleotide of the invention, for example, a coding sequence, or an antisense sequence. The construct can further comprises a regulatory element operably linked to the polynucleotide, e.g., an inducible regulatory element. The plant can be a dicotyledonous plant, e.g., a member of the *Solanaceae* family such as *Nicotiana tabacum*. The plant can also be a monocotyledonous plant, a gymnosperm, or a conifer.

A transgenic plant is disclosed herein that contains a polynucleotide expressing a polypeptide having from about 250 to about 550 amino acids. The polypeptide comprises an amino acid sequence substantially identical to the amino acid sequence of Figure 3.

A method of using a polynucleotide is disclosed herein. The method comprises the step of hybridizing the polynucleotide discussed above to DNA or RNA from a plant. The method can further comprise the steps of identifying a segment of the plant DNA or RNA that has about 70% or greater sequence identity to the

polynucleotide, and the step of cloning at least a portion of the DNA or RNA segment. The cloned portion may further comprise DNA flanking the segment having 70% or greater sequence identity.

In another aspect, the invention features a method of altering disease resistance in a plant. The method comprises the steps of introducing a polynucleotide of the invention into a plant cell; and producing a plant containing the polynucleotide from the plant cell. Expression of the polynucleotide alters disease resistance in the plant. For example, the nucleic acid construct may further comprise an inducible regulatory element operably linked to the polynucleotide and expression may be induced by the regulatory element upon exposure of the plant to an elicitor or plant pathogen.

In another aspect, the invention features an isolated polypeptide, having from about 250 to about 550 amino acids and comprising an amino acid sequence substantially identical to Figure 3.

An inducible regulatory element is a DNA sequence effective for regulating the expression of a polynucleotide that is operably linked to that regulatory element. For example, a CDPK gene product associated with a plant defense response (e.g., a hypersensitive response) can be operably linked to a developmentally-regulated regulatory element. Also included in this term are regulatory elements that are sufficient to render gene expression inducible in response to disease-associated external signals or agents (e.g., pathogen- or elicitor-induced signals or agents as described herein). Also included in this term are those regulatory elements flanking a novel CDPK gene and involved in rapid induction of transcription of such a novel gene. In general, defense response regulatory elements are located

5′ to the coding region of a gene, but are not so limited.

By "tissue-specific" is meant capable of preferentially increasing expression of a gene product (e.g., an mRNA molecule or polypeptide) in one tissue (e.g., xylem tissue) as compared to another tissue (e.g., phloem). By "cell-specific" is meant capable of preferentially increasing expression of a gene product (e.g., an mRNA molecule or polypeptide) in one cell (e.g., a parenchyma cell) as compared to another cell (e.g., an epidermal cell).

A "pathogen" is an organism whose infection of, or association with, cells of viable plant tissue can result in a disease. An "elicitor" is any molecule that is capable of initiating a plant defense response. Examples of elicitors include, without limitation, one or more toxic ions, e.g., mercuric ions, other chemically defined compositions, metabolic inhibitors, cell wall glycans, certain glycoproteins, certain enzymes, fungal spores, chitosans, certain fatty acids, and certain oligosaccharides derived from plant cell walls, and elicitins (e.g., harpin, cryptogein, and parasiticein).

By "operably linked" is meant that two polynucleotides are connected in such a way as to permit the two polynucleotides to achieve a desired functional activity, for example, linking of an inducible regulatory sequence and a coding sequence to achieve gene expression when the appropriate inducer molecules are present.

Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, suitable methods and materials are described

below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Brief Description of the Drawings

Figure 1 is a representation of the nucleotide sequences of the primers FokinB and RecalIV.

Figure 2 is a representation of the DNA sequence (SEQ ID NO:1) of a partial cDNA clone isolated from a cell suspension culture derived from a tobacco cultivar KY14 explant, after growth in the presence of the elicitin parasiticein.

Figure 3 is a representation of the deduced amino acid sequence of the DNA sequence of Figure 2, using the standard one letter amino acid code.

Figure 4 is a schematic comparison of the amino acid sequence of Figure 3 to that of a soybean CDPK.

Detailed Description of the Invention

The present invention relates to isolated polynucleotides (nucleic acids) that are induced in plant cells in response to invasion by a potential plant pathogen and/or treatment with an elicitor or elicitor-mimicking chemical signals. Such nucleic acids typically encode a calcium dependent protein kinase (CDPK) polypeptide or CDPK-related polypeptide. Induction of the novel polynucleotides disclosed herein corresponds in time to that of plant defense response genes, whereas

other CDPK genes appear to be induced less rapidly. Induction of gene expression for such novel genes is more rapid than that of genes involved in developmentally regulated processes in plants, e.g., developmentally regulated processes such as floral development. Induction of the novel CDPK genes disclosed herein is also more rapid than that of many genes involved in responses to abiotic stress, such as salt or water stress.

A polynucleotide of the present invention can be in the form of RNA or in the form of DNA, including cDNA, synthetic DNA or genomic DNA. The DNA can be double-stranded or single-stranded and, if single-stranded, can be either a coding strand or non-coding strand. An RNA analog of SEQ ID NO:1 may be, for example, mRNA or a combination of ribo- and deoxyribonucleotides.

A polynucleotide of the invention can encode a polypeptide including an amino acid sequence substantially similar or identical to that of Figure 3. In some embodiments, a polynucleotide may be a variant of the nucleic acid shown in SEQ ID NO:1, e.g., can have a different nucleotide sequence that, due to the degeneracy of the genetic code, encodes the same amino acid sequence as the polypeptide of Figure 3.

A polynucleotide of the invention can further include additional nucleic acid sequences. For example, a nucleic acid fragment encoding a secretory or leader amino acid sequence can be fused in-frame to the amino terminal end of a polypeptide comprising the amino acid sequence of Figure 3. Other nucleic acid fragments are known in the art that encode amino acid sequences useful for fusing in-frame to the CDPK polypeptides disclosed herein. See, e.g., U.S. 5,629,193. A polynucleotide can further include one or more regulatory elements operably linked to a CDPK polynucleotide disclosed herein.

The present invention also includes polynucleotides that selectively hybridize to a CDPK polynucleotide sequence disclosed herein. Hybridization may involve Southern analysis (Southern blotting), a method by which the presence of DNA sequences in a target nucleic acid mixture are identified by hybridization to a labeled oligonucleotide or DNA fragment probe. Southern analysis typically involves electrophoretic separation of DNA digests on agarose gels, denaturation of the DNA after electrophoretic separation, and transfer of the DNA to nitrocellulose, nylon, or another suitable membrane support for analysis with a radiolabeled, biotinylated, or enzyme-labeled probe as described in sections 9.37-9.52 of Sambrook et al., (1989) *Molecular Cloning*, second edition, Cold Spring Harbor Laboratory, Plainview, NY.

A polynucleotide can hybridize under moderate stringency conditions or under high stringency conditions to a polynucleotide disclosed herein. High stringency conditions are used to identify nucleic acids that have a high degree of homology or sequence identity to the probe. High stringency conditions can include the use of a denaturing agent such as formamide during hybridization, e.g., 50% formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM NaCl, and 75 mM sodium citrate at 42°C. Another example is the use of 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 μg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC and 0.1% SDS. Alternatively, low ionic strength and high temperature can be employed for washing, for example, 0.015 M NaCl/0.0015 M sodium citrate (0.1X SSC); 0.1% sodium lauryl sulfate (SDS) at 65°C.

Moderate stringency conditions are hybridization conditions used to identify nucleic acids that have less homology or identity to the probe than do nucleic acids identified under high stringency conditions. Moderate stringency conditions can include the use of higher ionic strength and/or lower temperatures for washing of the hybridization membrane, compared to the ionic strength and temperatures used for high stringency hybridization. For example, a wash solution comprising 0.060 M NaCl/0.0060 M sodium citrate (4X SSC) and 0.1% sodium lauryl sulfate (SDS) can be used at 50°C, with a last wash in 1X SSC, at 65°C. Alternatively, a hybridization wash in 1X SSC at 37°C can be used.

Hybridization can also be done by Northern analysis (Northern blotting), a method used to identify RNAs that hybridize to a probe. The probe is labeled with a radioisotope such as ^{32}P, by biotinylation or with an enzyme. The RNA to be analyzed can be electrophoretically separated on an agarose or polyacrylamide gel, transferred to nitrocellulose, nylon, or other suitable membrane, and hybridized with the probe, using standard techniques well known in the art such as those described in sections 7.39-7.52 of Sambrook et al., *supra*.

It is generally preferred that a probe of at least about 20 nucleotides in length be used, preferably at least about 50 nucleotides, more preferably at least about 100 nucleotides. If a relatively short probe is to be used, the nucleotide sequence of the probe preferably avoids regions conserved among plant CDPK genes (protein kinase domains and calcium-binding domains), to more readily distinguish the rapidly induced CDPK genes disclosed herein from more slowly induced CDPK genes, constitutive CDPK genes or low-level constitutive CDPK genes. Nevertheless, probes containing such conserved

regions can be used, provided that there are sufficient non-conserved regions present in the probe that are more specific for the novel polynucleotides disclosed herein.

A polynucleotide of the invention has at least about 70% sequence identity, preferably at least about 80% sequence identity, more preferably at least about 90% sequence identity to SEQ ID NO:1. Sequence identity can be determined, for example, by computer programs designed to perform single and multiple sequence alignments. Polynucleotides having at least about 70% nucleotide sequence identity to the polynucleotide of SEQ ID NO:1 are included in the invention and can be identified by hybridization under conditions of moderate stringency. Polynucleotides having at least about 80% sequence identity, or at least about 90% sequence identity, or at least about 95% sequence identity to the polynucleotide of SEQ ID NO:1 can be identified by high stringency hybridization.

A polynucleotide of the invention can be obtained by chemical synthesis, isolation and cloning from plant genomic DNA, or other means known to the art, including the use of polymerase chain reaction (PCR) technology carried out using oligonucleotides corresponding to portions of SEQ ID NO:1. PCR refers to a procedure or technique in which target nucleic acid is amplified in a manner similar to that described in U.S. Patent No. 4,683,195, incorporated herein by reference, and subsequent modifications of the procedure described therein. Generally, sequence information from the ends of the region of interest or beyond are employed to design oligonucleotide primers that are identical or similar in sequence to opposite strands of the template to be amplified. PCR can be used to amplify specific RNA sequences, specific DNA sequences from total genomic DNA, and cDNA transcribed from total cellular RNA,

bacteriophage or plasmid sequences, and the like. Alternatively, it is contemplated that a cDNA library (in an expression vector) can be screened with CDPK-specific antibody prepared using peptide sequence(s) from hydrophilic regions of the CDPK sequence of Figure 3 and technology known in the art.

The novel polynucleotides of the invention can be found in substantially all plants, including members of the *Leguminaceae* (e.g., soybean), members of the *Solanaceae* (e.g., *N. tabacum*), members of the *Brassicaceae* family (e.g., *Arabidopsis thaliana*) and members of the *Graminaceae* (e.g., *Zea mays*). Preferably, polynucleotides of the invention are selected from the *Solanaceae* family.

In some embodiments, a polynucleotide of the invention is identified and isolated from a plant based on nucleotide sequence homology and on the rapid induction of expression after elicitor or pathogen treatment. For example, DNA:DNA hybridization under conditions of moderate to high stringency with a polynucleotide probe disclosed herein allows the identification of corresponding genes from other plant species. Use of a target nucleic acid (e.g., cDNA) prepared from a tissue shortly after induction of defense responses facilitates the isolation of the novel polynucleotides disclosed herein, because such polynucleotides typically are more rapidly induced than other CDPK genes.

A nucleic acid construct comprises a polynucleotide as disclosed herein, and typically is linked to another, different polynucleotide. For example, a full-length CDPK coding sequence can be operably fused in-frame to a nucleic acid fragment that encodes a leader sequence, secretory sequence or other

additional amino acid sequences that may be usefully linked to a polypeptide or peptide fragment.

In some embodiments, a nucleic acid construct includes a polynucleotide of the invention operably linked to at least one suitable regulatory sequence in sense or antisense orientation. Regulatory sequences typically do not themselves code for a gene product. Instead, regulatory sequences affect the expression level of the coding sequence. Examples of regulatory sequences are known in the art and include, without limitation, minimal promoters and promoters of genes induced in response to elicitors. Native regulatory sequences of the polynucleotides disclosed herein can be readily isolated by those skilled in the art and used in constructs of the invention. Other examples of suitable regulatory sequences include enhancers or enhancer-like elements, introns, 3' non-coding regions such as poly A sequences and other regulatory sequences discussed herein. Molecular biology techniques for preparing such chimeric genes are known in the art.

Polypeptides of the invention have from about 250 to about 550 amino acids, e.g., from about 300 amino acids to about 508 amino acids, or from about 308 amino acids to about 500 amino acids. A polypeptide of the invention typically contains protein kinase domains as well as calcium-binding site domains. Such domains include, for example, amino acids 2 to 7, 42 to 49, 191 to 202, 227 to 238, 264 to 274, and 297 to 307 of Figure 3.

The amino acid sequence of the polypeptide can include the deduced amino acid sequence of Fig. 3. In other embodiments, a polypeptide of the invention includes an amino acid sequence substantially identical to that of Fig. 3, e.g., about 80% or greater sequence identity, or about 90% or greater sequence identity, or

about 95% or greater sequence identity. Generally, conservative amino acid substitutions or substitutions of similar amino acids are tolerated without affecting protein function. Similar amino acids are those that are similar in size and/or charge properties. For example, isoleucine and valine are similar amino acids. Similarity between amino acid pairs has been assessed in the art in a number of ways. For example, Dayhoff et al. (1978) in *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3, pp. 345-352, provides frequency tables for amino acid substitutions which can be employed as a measure of amino acid similarity. Protein kinase domains and calcium-binding site domains may be altered by conservative substitutions, but generally are retained without alterations in amino acid sequence.

An "isolated" polypeptide is expressed and produced in a manner or environment other than the manner or environment in which the polypeptide is naturally expressed and produced. For example, a polypeptide is isolated when expressed and produced in bacteria or fungi. Similarly, a polypeptide is isolated when a gene encoding it is operably linked to a chimeric regulatory element and expressed in a tissue or species where the polypeptide is not naturally expressed. In addition, a polypeptide is isolated when a gene encoding it is operably linked to a chimeric regulatory element and is expressed in a tissue where the polypeptide is naturally expressed, but at higher levels. A polypeptide of the invention can also be isolated by standard purification methods to obtain it in about 80% or greater purity, or about 90% or greater purity or about 95% or greater purity.

In some embodiments, a polypeptide of the invention is an analog or variant of a polypeptide including the deduced amino acid sequence of Fig. 3.

Such analogs or variants include, for example, naturally occurring allelic variants, non-naturally occurring allelic variants, deletion variants, and insertion variants, that do not substantially alter the function of the polypeptide.

A polypeptide of the invention may comprise the sequence shown in Fig. 3 as well as the flanking amino terminal and carboxy terminal sequences encoded by the same gene as that comprising the nucleotide sequence of SEQ ID NO:1. Alternatively, a chimeric polypeptide may be produced from a gene that links, in-frame, nucleotides from the 5′ region of a first CDPK gene to nucleotides from the 3′ region of a second CDPK gene, thereby forming a chimeric gene that encodes the chimeric polypeptide. An illustrative example of a chimeric CDPK polypeptide is a polypeptide expressed by a polynucleotide encoding amino acids 1 to 156 from the amino terminal region of a soybean CDPK gene (Fig. 4), followed by the amino acid sequence of Fig. 3, followed by amino acids 465 to 508 from the carboxy terminal region of the same soybean CDPK gene, all of which are fused in-frame.

A transgenic plant of the invention contains a nucleic acid construct as described herein. Such a construct is introduced into a plant cell and at least one transgenic plant is obtained. Seeds produced by a transgenic plant can be grown and selfed (or outcrossed and selfed) to obtain plants homozygous for the construct. Seeds can be analyzed to identify those homozygotes having the desired expression of the construct. Transgenic plants may be entered into a breeding program, e.g., to increase seed, to introgress the novel construct into other lines or species, or for further selection of other desirable traits. Alternatively, transgenic plants may be obtained by

vegetative propagation of a transformed plant cell, for those species amenable to such techniques.

As used herein, a transgenic plant also refers to progeny of an initial transgenic plant. Progeny includes descendants of a particular plant or plant line, e.g., seeds developed on an instant plant. Progeny of an instant plant also includes seeds formed on F_1, F_2, F_3, and subsequent generation plants, or seeds formed on BC_1, BC_2, BC_3, and subsequent generation plants.

In some embodiments, a transgenic plant contains a construct that includes a polynucleotide of the invention operably linked in sense orientation to a suitable regulatory element, so that a sense mRNA is produced. If desired, a selectable marker gene can be incorporated into the construct in order to facilitate identification of transformed cells or tissues.

Inhibition of the novel CDPK genes in plants is also useful. For example, inhibition of CDPK gene expression shortly before harvest of a seed crop can permit plant pathogens to more readily invade plant vegetative tissues, thereby reducing the amount of plant biomass that interferes with mechanical harvesting of the seeds. Regulated inhibition of CDPK gene expression can be accomplished by operably linking, in antisense orientation, a polynucleotide of the invention to a suitable inducible regulatory sequence. See, e.g., U.S. Patent 5,453,566. One can achieve the same effect by cosuppression, i.e, expression in the sense orientation of the entire or partial coding sequence of a novel CDPK gene can suppress corresponding endogenous CDPK genes. See, e.g., WO 94/11516.

In some embodiments, a nucleic acid construct includes a polynucleotide disclosed herein, operably linked to a minimal promoter. Such a construct, when introduced into and expressed in a plant, can confer low

level constitutive expression of the polynucleotide, resulting in an enhanced systemic defense response by the plant. A minimal promoter contains the DNA sequence signals necessary for RNA polymerase binding and initiation of transcription. Generally, transcription directed by a minimal promoter is low and does not respond either positively or negatively to environmental or developmental signals in plant tissue. An exemplary minimal promoter suitable for use in plants is the truncated CaMV 35S promoter, which contains the region from -90 to +8 of the 35S transcription unit.

Transcriptional regulatory sequences can be used to control gene expression in suspension cultures. For example, the EAS4 promoter including the transcription initiation signals, the inducible transcription regulatory element and the transcription-enhancing element, can be used to mediate the inducible expression of the disclosed coding sequence in transgenic plants or suspension cell cultures. See U.S. Application Serial No. 08/577,483. When desired, expression of the coding sequence of interest is induced by the application of an elicitor or other inducing signal.

Transgenic techniques for use in the invention include, without limitation, *Agrobacterium*-mediated transformation, electroporation and particle gun transformation. Illustrative examples of transformation techniques are described in U.S. Patent 5,204,253, (particle gun) and U.S. Patent 5,188,958 (*Agrobacterium*). Transformation methods utilizing the Ti and Ri plasmids of *Agrobacterium spp*. typically use binary type vectors. Walkerpeach, C. et al., in Plant Molecular Biology Manual, S. Gelvin and R. Schilperoort, eds., Kluwer Dordrecht, C1:1-19 (1994).

In some embodiments, an inducible transcription regulatory sequence can be coupled to a promoter sequence

functional in plants, both of which are operably linked to a polynucleotide of the invention. When such a regulatory element is coupled to a promoter, a truncated (or minimal) promoter generally is used, for example, the truncated 35S promoter of Cauliflower Mosaic Virus (CaMV). Truncated versions of other constitutive promoters can also be used, e.g., *A. tumefaciens* T-DNA genes such as *nos*, *ocs*, and *mas*, and plant virus genes such as the CaMV 19S gene.

Techniques are well-known to the art for the introduction of DNA into monocots as well as dicots, as are the techniques for culturing plant tissues and regenerating those tissues. Monocots which have been successfully transformed and regenerated include wheat, corn, rye, rice and asparagus. See, e.g., U.S. Patent Nos. 5,484,956 and 5,550,318. Transgenic aspen tissue has been prepared and transgenic plants have been regenerated. Poplars have also been transformed. Technology is also available for the manipulation, transformation, and regeneration of Gymnosperm plants. See, e.g., U.S. Patent No. 5,122,466 and U.S. Patent No. 5,041,382.

A method according to the invention includes the introduction of a nucleic acid construct into a plant cell and the production of a plant from such a transformed cell. Expression of the polynucleotide present in the construct alters the disease resistance phenotype of the plant, e.g., a novel disease resistance phenotype is conferred on the plant or an existing disease resistance phenotype is enhanced.

Disease resistance phenotype involves the level and timing of host defensive responses in the transgenic plant. Assays to indicate that disease resistance has been altered typically include the application of a compound that ordinarily elicits a defensive response to

a transgenic plant and, in parallel, the application of the same compound to a control plant. A control plant typically is from the same parental line as the one into which a new nucleic acid construct was introduced. Disease resistance is enhanced or conferred on a plant by expression of a polynucleotide of the invention when there is a higher level of resistance in the transgenic plant than the corresponding resistance in the control plant. Disease resistance can be measured with reference to a specific pathogen, e.g., a *Phythophthora spp.*. Disease resistance can also be measured with reference to several pathogens, to identify an enhanced systemic defense response.

Where transgenic plants are to be induced for expression of a CDPK coding sequence operably linked to an elicitor-mediated regulatory element, the elicitor typically must penetrate the cuticle of the plant to have an inductive effect. Plant tissue can be wounded to facilitate or allow the uptake of the elicitor into the plant tissue. A wide variety of inducing compositions, including elicitors and other chemical signals, such as the combination of ethylene and methyl jasmonate, can be effectively used to induce expression.

A method of using a polynucleotide of the invention comprises the step of hybridizing the polynucleotide to DNA or RNA from a plant. Hybridization can be carried out, for example, as described hereinabove. The method can further comprise the step of identifying a segment of the plant DNA or RNA that has a significant degree of sequence identity to the polynucleotide, e.g., 70% sequence identity, preferably 80% sequence identity, 90% sequence identity, or 95% sequence identity. The segment can be identified by electrophoretic separation of the plant DNA or RNA and the use of labeled polynucleotide probe, which results in

a visible band at the position of the homologous segment. Segments can be generated, for example, by physical shearing or by restriction endonuclease digestion. A segment can be as short as 100 bp (nucleotides) in length, but typical segments are at least 1000 bp, and can be 10,000 bp or greater.

Such a method can further comprise the step of cloning at least a portion of the DNA or RNA segment, including, but not limited to, DNA flanking the homologous segment. Such flanking DNA can include promoters, enhancers, transcriptional regulatory elements and poly A sequences. Flanking DNA can be either 5′ to or 3′ to the homologous segment and preferably includes 300, or 600, or 1,000 bp of DNA beyond the coding sequence, because regulatory elements generally are found within this span.

Promoters and other elicitor or pathogen-responsive regulatory elements flanking the novel polynucleotides disclosed herein are particularly useful, because such elements confer very rapid induction of gene expression after treatment with pathogen or elicitor. Such regulatory elements can be operably linked to useful genes to allow rapid production of desirable compounds. For example, such regulatory elements can be used to drive expression of genes encoding antibodies, blood clotting factors, antigenic peptides, viral replicases or coat proteins, and enzymes involved in secondary metabolite synthesis (such as isoprenoid biosynthesis). See, e.g., U.S. Patent 5,612,487; U.S. Patent 5,484,719; and U.S. Application Ser. No. 08/577,483, filed December 22, 1995.

After introducing a chimeric gene having an elicitor or pathogen-responsive element into a plant, expression of the chimeric gene product can be induced with an appropriate pathogen or elicitor. Production of

the desired gene product (or its enzymatic end product) rapidly ensues and the desired product can then be obtained.

The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

EXAMPLES

The following examples use many techniques well-known and accessible to those skilled in the arts of molecular biology, in the manipulation of recombinant DNA in plant tissue and in the culture and regeneration of transgenic plants. Enzymes are obtained from commercial sources and are used according to the vendors' recommendations or other variations known to the art. Reagents, buffers, and culture conditions are also known to the art. Abbreviations and nomenclature, where employed, are deemed standard in the field and are commonly used in professional journals such as those cited herein.

Example 1.

Cloning of a Tobacco CDPK cDNA

The elicitor parasiticein was prepared by expression of the *Phytophthora parA1* gene in *E. coli* cells and isolation of the gene product from the periplasmic space.

Genomic DNA of *Phytophthora* Race 0 was isolated from mycelium essentially as described in Xu, J., et al. Trends in Genetics 10:226-227 (1994). The DNA was sheared and used as a template for PCR amplification of the *parA1* gene, using primers designed according to the *parA1* sequence reported in Kamoun, S., et al. Mol. Plant-Microbe Interact. 6:573-581 (1993). The *parA1* PCR product was cloned into pBluescript (Stratagene, San

Diego, CA) and the sequence of the product determined by double-stranded DNA sequencing using the dideoxy chain termination method.

The *parA1* insert in pBluescript was amplified by PCR, using primers that created an N-terminal histidine tag and a protein kinase site at the 5′ end of the gene. The PCR product was ligated into the expression vector pET28b (Novagen, Madison, WI) and, after confirming the DNA sequence of the *parA1* fusion, the pET28b construct was transformed into *E. coli* BL21.

A BL21 culture containing the *parA1* fusion was grown at 37° C in the presence of kanamycin to an OD_{600} of 0.3. IPTG (1mM) was added and the culture was incubated for 5 hours at 27° C.

Periplasmic proteins were prepared by osmotic shock essentially as described in Ausubel, F., et al. in Current Protocols in Molecular Biology, John Wiley & Sons, New York (1989). Cells (1.5 ml) were harvested by centrifugation, resuspended in 500 μl of 50 mM Tris-HCl, pH 8.0, 20% sucrose, 1 mM EDTA and incubated with shaking for 10 minutes at room temperature. After centrifugation, the pellet was resuspended in 200 μl ice cold $MgSO_4$ (5 mM) and incubated with shaking for 10 minutes at 4° C. The mixture was centrifuged and the resulting supernatant (containing periplasmic proteins) was applied to a Ni^{++} column. The *parA1* protein was purified from the column according to the manufacturer's directions. The protein concentration in the *parA1* extract was determined by the Bradford method.

Nicotiana tabacum L. cv. KY14 cell suspension cultures were treated with parasiticein at a final concentration of 2 μg/ml during rapid growth phase to induce stress response genes. Parallel suspension cell cultures which were not treated with parasiticein served as controls. Cells were collected by gentle vacuum

filtration 0, 30, 60 and 120 minutes after the addition of elicitor.

Total RNA was isolated from treated and untreated tobacco cells and used as template for targeted differential display reverse transcriptase PCR (TDDRT-PCR). First strand cDNA was generated using a cDNA cycle kit from Invitrogen (San Diego, CA). The first strand cDNAs were then used as templates for PCR. The PCR reaction was carried out using typical conditions as described in PCR Protocols: A Guide to Methods and Applications, Innis, M., Gelfand, D., Sminsky, J. and White, T., eds. Academic Press Inc., San Diego, CA (1990), except that the annealing temperature was 58°C. The PCR primers were FokinB (GTTGACTCCCTACCCTCTT) and RecalIV (GGTACTTAGGAAGTGTTACGGG). See Figure 1. PCR products were separated by electrophoresis on a 1% (w/v) agarose gel and products of greater than about 800 base pairs (bp) from the 60 minute treated culture were purified by electroelution onto DE-81 paper (Whatman). Ends of the purified PCR products were filled in with Klenow polymerase, ligated to the EcoRV site of pBluescript, and transformed into *E. coli* TB1.

Ampicillin resistant TB1 colonies were screened for the presence of a $\geq$800 bp DNA fragment inserted into pBluescript. The sequence of one such insert was determined by the dideoxynucleotide chain termination procedure of Sanger et al. (1977) *Proc. Natl. Acad. Sci. USA* 74:8073-8077, with a Sequenase® kit from United States Biochemical Corp., Cleveland, OH) or an automated fluorescence based system (Applied Biosystems, Foster City, CA). The sequence of the insert in the vector was determined on both strands. The plasmid containing this insert was designated pCDPK-1.

The nucleotide sequence of the insert in pCDPK-1 is shown in Figure 2 and the deduced amino acid sequence

of the insert is shown in Figure 3. The deduced amino acid sequence was compared to amino acid sequences of plant genes in the GenBank, EMBL, and Swiss Prot databases. Homology was found to plant CDPK polypeptides, including polypeptides from *Glycine max*, *Arabidopsis thaliana*, *Vigna radiata*, *Zea mays* and *Cucurbita pepo*.

Using the BLASTP program and a BLOSUM62 scoring matrix, two regions of homology to serine/threonine protein kinase domains were identified in the amino terminal portion of the polypeptide and four regions of homology to Ca^{++} binding domains were identified in the carboxyl terminal portion of the polypeptide. Figure 4 shows a comparison of the amino acid sequence of Fig. 3 and a soybean CDPK amino acid sequence (Genbank Accession No:M64987). The amino acid sequence of the tobacco calcium binding sites were similar to the amino acid sequence of corresponding sites in the soybean CDPK. However, there were significant differences in other parts of the sequence. The comparison indicates that there is about 78% overall sequence identity between the soybean CDPK and CDPK-1.

The BLASTN program was used to compare the pCDPK-1 nucleotide sequence to nucleic acid sequences on various databases. Based on the nucleotide sequence of other plant CDPK genes and the length of the polypeptides encoded thereby, the nucleic acid insert present in pCDPK-1 is estimated to lack about 560 bp of 5′ CDPK-1 coding sequence and about 130 bp of 3′ CDPK-1 coding sequence.

Example 2.

Isolation of a full-length cDNA clone

To obtain a full-length clone, a RACE (Rapid Amplification of cDNA Ends) approach is used, with polyA+ RNA prepared from tobacco cells after induction with

elicitor being the template. PolyA+ RNA is prepared as described in Example 1.

A primer having the sequence GAC AAG GAC GGG AGT GGG TAT (Primer A, internal to CDPK-1) and a primer having the sequence GAC TCG AGT CGA CAT CGA TTT TTT TTT TTT TTT TT (dT_{17} adapter-primer) are used to amplify the 3' end of the CDPK coding sequence. The reverse transcriptase reaction is carried out in 2 μl 10X RTC buffer, 10 units of RNasin (Promega Biotech), 0.5 μg of dT_{17} adapter-primer and 10 Units of AMV reverse transcriptase (Life Sciences) in a total volume of 3.5 μl, as described in Frohman, M. in PCR Protocols: A Guide to Methods and Applications, *supra*, pp. 28-38. The PCR amplification reaction is carried out in 5 μl 10X PCR buffer, 5 μl DMSO, 5 μl 10X dNTPs (15 mM each), 30 μl H_2O, 1 μl adapter-primer (25 pmol, GAC TCG AGT CGA CAT CG), 1 μl primer A and 1-5μl cDNA. Cycle times are as indicated in Frohman, *supra*.

The 5' end of the CDPK coding sequence is cloned by carrying out reverse transcription as described above, using 10 pmole of primer B (AGG GGC TAC GTA GTA AGG ACT) instead of dT_{17} adapter-primer. The cDNA product is extended using terminal transferase and dATP as described in Frohman, *supra*, and then amplified by PCR as described above with 10 pmole of dT_{17} adapter-primer, 10 pmole of adapter-primer and 10 pmole of primer C (ATT CTC AGG CTT AAG GTC CCT). PCR is carried out under standard conditions. Back et al. (1994) *Arch. Biochem. Biophys.* 315:523-532. The amplified 3' and 5' products are blunt-end cloned into pBluescript SK (Stratagene) and combined with the pCDPK-1 insert by routine molecular biology techniques to form a full-length cDNA of the tobacco CDPK coding sequence.

The DNA sequence of the full-length cDNA is determined by a dideoxynucleotide chain termination procedure, as described in Example 1.

Example 3.

Induction of CDPK-Homologous RNA in Tobacco Suspension Cultures

The DNA insert in pCDPK-1 was used as a probe to follow the induction of gene expression in response to elicitor. *Nicotiana tabacum* L. cv. KY14 cell suspension cultures were treated with parasiticein for 0, ½, 1, 2, 6 and 12 hours as described in Example 1. Total RNA was isolated and electrophoresed on a 1% agarose gel. The insert from pCDPK-1 was radiolabeled by the random priming method and hybridized to the gel-separated RNA as described in Sambrook, J. et al., *supra*. No mRNA hybridizing to CDPK-1 was detected prior to elicitor treatment, whereas mRNA hybridizing to CDPK-1 was readily detected at 1/2, 1 and 2 hours after elicitor treatment. At 6 and 12 hours after elicitor treatment, no mRNA hybridizing to CDPK-1 could be detected, indicating that CDPK-1 gene expression had decreased to undetectable levels by about 6 hours.

Example 4.

Construction of a Chimeric CDPK Gene

A CDPK gene is constructed from: a chemically synthesized DNA encoding amino acids 1 to 156 of the soybean CDPK of Figure 6, a chemically synthesized DNA encoding amino acids 465 to 508 of the soybean CDPK of Figure 6, and the CDPK insert of pCDPK-1. The three DNAs are ligated by routine molecular biology techniques to form a chimeric CDPK coding sequence having amino acids 1 to 156 of soybean CDPK at the amino terminal end, fused

in-frame to amino acids 1 to 307 of tobacco CDPK (Fig. 3), which in turn is fused in-frame to amino acids 465 to 508 of soybean CDPK at the carboxyl terminal end.

The chimeric coding sequence is inserted in sense orientation into an *Agrobacterium* binary vector containing a minimal 35S and EAS4 inducible regulatory element. Operable linkage of the regulatory element, promoter, and coding sequence is confirmed by determining the DNA sequence of the junction regions and by expression in transgenic plants.

Example 5.

Generation of Transgenic Plants

Transformed plant cell lines are produced using a modified *Agrobacterium tumefaciens* transformation protocol. Nucleic acid constructs are prepared that contain the full-length CDPK cDNA of Example 3 or the chimeric CDPK coding sequence of Example 4. The recombinant constructs containing the sequences to be introduced into plants are transferred into *A. tumefaciens* strain GV3850 by triparental mating with *E. coli* TB1 (pRK2013). *N. tabacum* leaves at a variety of stages of growth are cut into 1 cm^2 pieces, and dipped in a suspension of Agrobacterium cells (about 10^4 to 10^5 cells/ml). After 3 to 10 minutes, the leaf segments are then washed in sterile water to remove excess bacterial cells and to reduce problems with excess bacterial growth on the treated leaf segments. After a short drying time (30 to 60 seconds), the treated leaf segments are placed on the surface of Plant Tissue Culture Medium without antibiotics to promote tissue infection and DNA transfer from the bacteria to the plant tissue. Plant Tissue Culture Medium contains per liter: 4.31 g of Murashige and Skoog Basal Salts Mixture (Sigma Chemical Company, St. Louis, MO), 2.5 mg of benzylaminopurine (dissolved in

1 N NaOH), 10 ml of 0.1 mg/ml indoleacetic acid solution, 30 g sucrose, 2 ml of Gamborg's Vitamin Solution (Sigma Chemical Co., St. Louis, MO) and 8 g of agar. The pH is adjusted between pH 5.5 and 5.9 with NaOH. After 2 days, the leaf segments are transferred to Plant Tissue Culture Medium containing 300 μg/ml of kanamycin, 500 μg/ml of mefoxin (Merck, Rahway, NJ). Kanamycin selects for transformed plant tissue, and mefoxin selects against *Agrobacterium*.

It may be necessary to minimize the exposure of the explant tissue to *Agrobacterium* cells during the transformation procedure if a pathogen-inducible regulating element is used, because *Agrobacterium* cells may themselves induce the element after introduction into the plant cells. Accordingly, the biolistic technique for the introduction of DNA containing cell suicide genes under the regulatory control of the inducible transcriptional regulatory element is a useful alternative transformation technique because it does not entail the use of Agrobacterium cells or fungal cell wall digestive enzymes (as necessary for the generation of protoplasts for electroporation), both of which can lead to induction of the coding sequences under the control of that regulatory element.

Transgenic plants are regenerated essentially as described by Horsch et al. (1985) *Science* 227:1229-1231.

Example 6.

Elicitor- and Pathogen-inducible Expression of a Chimeric CDPK Gene in Transgenic Tobacco

The activity of the CDPK constructs of Example 7 are measured in transgenic tobacco plants treated with either an elicitor or pathogen. As controls, transgenic tobacco plants expressing the GUS reporter gene under the control of the cauliflower mosaic virus (CaMV) 35S

promoter are also produced. F_1 seeds from regenerated transgenic tobacco plants are germinated on medium containing 100 mg/L kanamycin. The resulting kanamycin-resistant plants are subsequently transferred into soil and grown in a greenhouse. Half of the plants are tested for the expression of the CDPK gene under inducing conditions, e.g., by intercellular application of elicitor or cellulase to the transgenic plants. Elicitor or cellulase is applied with a mechanical pipetter. As a control, remaining plants are mock-treated with a solution lacking cellulase or elicitor. Tobacco tissue is wounded with a scalpel in some experiments to facilitate exposure to the inducing compound.

Example 7.

Identification of CDPK Homologous Sequences

Tobacco leaf genomic DNA is isolated as described in Murray and Thompson (1980) *Nucleic Acids Research* 8:4321-4325. After digestion of aliquots with desired restriction enzymes, the digested DNA samples are electrophoresed on 0.8% agarose gels and the size-separated DNAs are transferred to nylon membranes. DNA blots are hybridized with the 900 bp CDPK cDNA insert of Example 1 that is radiolabeled by the random primer method. Hybridization is performed at 60°C in 0.25 M sodium phosphate buffer, pH 8.0, 0.7% SDS, 1% bovine serum albumin, 1 mM EDTA. The blot is then washed twice at 45°C with 2X SSC, 0.1% SDS and twice with 0.2X SSC, 0.1% SDS (1X SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0). Relative hybridization intensities of the various bands on the membrane are estimated from autoradiograms using a video densitometer (MilliGen/Biosearch, Ann Arbor, MI).

To identify polynucleotides having homologous sequences to tobacco CDPK and to determine the apparent number of copies per genome of those sequences, Southern hybridization experiments are carried out using target DNA isolated from other plant species and tobacco CDPK probes. Restriction endonuclease-digested genomic DNAs of various plant species are separated by agarose gel electrophoresis (0.8% agarose), and then transferred to a Hybond-N^+ membrane (Amersham Corp., Arlington Heights, IL). Radiolabeled probe fragments comprising coding sequences of pCDPK-1 are hybridized to the digested genomic DNA essentially as described in Sambrook et al. (1989), <u>supra</u>. Moderate stringency conditions are used (hybridization in 4X SSC, at 65°C with the last wash in 1X SSC, at 65°C).

Alternatively, PCR is carried out using target genomic DNA as a template and primers derived from highly conserved regions of the pCDPK-1 coding sequence.

Example 8.

Genomic DNA Flanking a CDPK Coding Sequence

The cDNA clone described in Example 1 is used as a hybridization probe for screening a *N. tabacum* cv. NK326 genomic library in the λEMBL3 vector (Clontech, Palo Alto, CA). Genomic DNA clones having 70% or greater sequence identity to the tobacco CDPK of Example 1 are identified using routine subcloning protocols. The nucleotide sequences of the cloned nucleic acid inserts are determined using routine DNA sequencing protocols.

One of the genomic DNA clones has a full-length coding sequence that comprises the tobacco CDPK coding sequence of Example 1. The clone also contains DNA contiguous with, and 5′ to, the coding sequence of Example 1. Examination of the nucleotide sequence of the 5′ flanking DNA in this clone reveals a putative ATG

start codon as well as one or more putative regulatory elements upstream of the start codon and within about 1000 bp of the start codon.

Other Embodiments

It is to be understood that while the invention has been described in conjunction with the Detailed Description thereof, that the foregoing description is intended to illustrate, and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: University of Kentucky Research Foundation

(ii) TITLE OF THE INVENTION: PROTEIN KINASES AND USES THEREOF

(iii) NUMBER OF SEQUENCES: 11

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C., P.A.
(B) STREET: 60 South Sixth Street, Suite 3300
(C) CITY: Minneapolis
(D) STATE: MN
(E) COUNTRY: USA
(F) ZIP: 55402

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 07-JUL-1998
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/889,655
(B) FILING DATE: 08-JUL-1997

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Lundquist, Ronald C
(B) REGISTRATION NUMBER: 37,875
(C) REFERENCE/DOCKET NUMBER: 07678/020WO1

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 612-335-5050
(B) TELEFAX: 612-288-9696
(C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 921 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
AGGGACCTTA AGCCTGAGAA TTTCCTTTTC AGTGCCGACG ACTTCATGGT AAAGAGTAAG  60
GCCACCGACT TCGGGCTTAG TGTATTCTAT AAGCCTGGGC AAAAGTTCAC GGACATAGTA 120
GGGAGTCCTT ACTACGTAGC CCCTGAGGTA CTTAGGAAGT GTTACGGGCC TGGGAGTGAC 180
GTATGGAGTG CCGGGGTAAT ACTTTACACC CTTCTTTGTG GGGCCCCTCC TTTCATGGCC 240
GACAGTGAGC CTGGGGTAGC CCTTCAAATA CTTCATGGGG ACCTTGACTT CAAGAGTGAC 300
CCTTGGCCTA CCATAAGTGA GAGTGCCAAG GACCTTATAA GGAAGATGCT TGAGCAAGAC 360
CCTAAGAGGA GGCTTACCGC CCATGAGGTA CTTAGGCATC CTTGGATAGT AGACGAGAAT 420
```

```
ATAGCCCCTG ACAAGCCTCT TGGGCCTGCC GTACTTAGTA GGCTTAAGCA ATTCAGTGCC 480
ATGAATAAGA TAAAGAAGAT GGCCCTTAGG GTAATAGCCG AGAGGCTTAG TGAGGAGGAG 540
ATAGTAGGGC TTAAGGAGAT GTTCAAGATG GACACCGACA ATAGTGGGAC CGTAACCTTC 600
TTCCATCTTA AGCAAGGGCT TAAGAGGGTA GGGAGTCAAC TTGGGGAGAG TGAGATAAAG 660
GACCTTATGG ACGCCGCCGA CGTAGACAAT AGTGGGACCA TAGACTATGG GGAGTTCGTA 720
ACCGCCGCCA TGCATCTTAA TAAGATAAAG AGGGAGGACC ATCTTGTAAG TGCCTTCAGT 780
TATCATGACA AGGACGGGAG TGGGTATATA GAGGTAGACG AGCTTAGGCA AGCCCTTGAG 840
GAGTTCGGGG,TACCTGACAC CAGTCTTGAG GACATGATAA AGGAGGTAGA CACCGACAAT 900
GATGGGCAAA TAGATTATGG G                                           921
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 307 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Arg Asp Leu Lys Pro Glu Asn Phe Leu Phe Ser Ala Asp Asp Phe Met
 1               5                  10                  15
Val Lys Ser Lys Ala Thr Asp Phe Gly Leu Ser Val Phe Tyr Lys Pro
            20                  25                  30
Gly Gln Lys Phe Thr Asp Ile Val Gly Ser Pro Tyr Tyr Val Ala Pro
        35                  40                  45
Glu Val Leu Arg Lys Cys Tyr Gly Pro Gly Ser Asp Val Trp Ser Ala
    50                  55                  60
Gly Val Ile Leu Tyr Thr Leu Leu Cys Gly Ala Pro Pro Phe Met Ala
65                  70                  75                  80
Asp Ser Glu Pro Gly Val Ala Leu Gln Ile Leu His Gly Asp Leu Asp
                85                  90                  95
Phe Lys Ser Asp Pro Trp Pro Thr Ile Ser Glu Ser Ala Lys Asp Leu
            100                 105                 110
Ile Arg Lys Met Leu Glu Gln Asp Pro Lys Arg Arg Leu Thr Ala His
        115                 120                 125
Glu Val Leu Arg His Pro Trp Ile Val Asp Glu Asn Ile Ala Pro Asp
    130                 135                 140
Lys Pro Leu Gly Pro Ala Val Leu Ser Arg Leu Lys Gln Phe Ser Ala
145                 150                 155                 160
Met Asn Lys Ile Lys Lys Met Ala Leu Arg Val Ile Ala Glu Arg Leu
                165                 170                 175
Ser Glu Glu Glu Ile Val Gly Leu Lys Glu Met Phe Lys Met Asp Thr
            180                 185                 190
Asp Asn Ser Gly Thr Val Thr Phe Phe His Leu Lys Gln Gly Leu Lys
        195                 200                 205
Arg Val Gly Ser Gln Leu Gly Glu Ser Glu Ile Lys Asp Leu Met Asp
    210                 215                 220
Ala Ala Asp Val Asp Asn Ser Gly Thr Ile Asp Tyr Gly Glu Phe Val
225                 230                 235                 240
Thr Ala Ala Met His Leu Asn Lys Ile Lys Arg Glu Asp His Leu Val
                245                 250                 255
Ser Ala Phe Ser Tyr His Asp Lys Asp Gly Ser Gly Tyr Ile Glu Val
            260                 265                 270
Asp Glu Leu Arg Gln Ala Leu Glu Glu Phe Gly Val Pro Asp Thr Ser
        275                 280                 285
Leu Glu Asp Met Ile Lys Glu Val Asp Thr Asp Asn Asp Gly Gln Ile
    290                 295                 300
Asp Tyr Gly
305
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGTACTTAGG AAGTGTTACG GG 22

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTTGACTCCC TACCCTCTT 19

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 512 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Ala Ala Lys Ser Ser Ser Ser Ser Thr Thr Thr Asn Val Val Thr
 1               5                   10                  15
Leu Lys Ala Ala Trp Val Leu Pro Gln Arg Thr Gln Asn Ile Arg Glu
             20                  25                  30
Val Tyr Glu Val Gly Arg Lys Leu Gly Gln Gly Gln Phe Gly Thr Thr
         35                  40                  45
Phe Glu Cys Thr Arg Arg Ala Ser Gly Gly Lys Phe Ala Cys Lys Ser
     50                  55                  60
Ile Pro Lys Arg Lys Leu Leu Cys Lys Glu Asp Tyr Glu Asp Val Trp
 65                  70                  75                  80
Arg Glu Ile Gln Ile Met His His Leu Ser Glu His Ala Asn Val Val
                 85                  90                  95
Arg Ile Glu Gly Thr Tyr Glu Asp Ser Thr Ala Val His Leu Val Met
             100                 105                 110
Glu Leu Cys Glu Gly Gly Glu Leu Phe Asp Arg Ile Val Gln Lys Gly
         115                 120                 125
His Tyr Ser Glu Arg Gln Ala Ala Arg Leu Ile Lys Thr Ile Val Glu
     130                 135                 140
Val Val Glu Ala Cys His Ser Leu Gly Val Met His Arg Asp Leu Lys
145                 150                 155                 160
Pro Glu Asn Phe Leu Phe Asp Thr Ile Asp Glu Asp Ala Lys Leu Lys
                 165                 170                 175
Ala Thr Asp Phe Gly Leu Ser Val Phe Tyr Lys Pro Gly Glu Ser Phe
             180                 185                 190
Cys Asp Val Val Gly Ser Pro Tyr Tyr Val Ala Pro Glu Val Leu Arg
         195                 200                 205
Lys Leu Tyr Gly Pro Glu Ser Asp Val Trp Ser Ala Gly Val Ile Leu
```

```
    210                 215                 220
Tyr Ile Leu Leu Ser Gly Val Pro Pro Phe Trp Ala Glu Ser Glu Pro
225                 230                 235                 240
Gly Ile Phe Arg Gln Ile Leu Leu Gly Lys Leu Asp Phe His Ser Glu
                245                 250                 255
Pro Trp Pro Ser Ile Ser Asp Ser Ala Lys Asp Leu Ile Arg Lys Met
            260                 265                 270
Leu Asp Gln Asn Pro Lys Thr Arg Leu Thr Ala His Glu Val Leu Arg
        275                 280                 285
His Pro Trp Ile Val Asp Asp Asn Ile Ala Pro Asp Lys Pro Leu Asp
    290                 295                 300
Ser Ala Val Leu Ser Arg Leu Lys Gln Phe Ser Ala Met Asn Lys Leu
305                 310                 315                 320
Lys Lys Met Ala Leu Arg Val Ile Ala Glu Arg Leu Ser Glu Glu Glu
                325                 330                 335
Ile Gly Gly Leu Lys Glu Leu Phe Lys Met Ile Asp Thr Asp Asn Ser
            340                 345                 350
Gly Thr Ile Thr Phe Asp Glu Leu Lys Asp Gly Leu Lys Asp Gly Leu
        355                 360                 365
Lys Arg Val Gly Ser Glu Leu Met Glu Ser Glu Ile Lys Asp Leu Met
    370                 375                 380
Asp Ala Ala Asp Ile Asp Lys Ser Gly Thr Ile Asp Tyr Gly Glu Phe
385                 390                 395                 400
Ile Ala Ala Thr Val His Leu Asn Lys Leu Glu Arg Glu Glu Asn Leu
                405                 410                 415
Val Ser Ala Phe Ser Tyr Phe Asp Lys Asp Gly Ser Gly Tyr Ile Thr
            420                 425                 430
Leu Asp Glu Ile Gln Gln Ala Cys Lys Asp Phe Gly Leu Asp Asp Ile
        435                 440                 445
His Ile Asp Asp Met Ile Lys Glu Ile Asp Gln Asp Asn Asp Gly Gln
    450                 455                 460
Ile Asp Tyr Gly Glu Phe Ala Ala Met Met Arg Lys Gly Asn Gly Gly
465                 470                 475                 480
Ile Gly Arg Arg Thr Met Arg Lys Thr Leu Asn Leu Arg Asp Ala Leu
                485                 490                 495
Gly Leu Val Asp Asn Gly Ser Asn Gln Val Ile Glu Gly Tyr Phe Lys
            500                 505                 510
```

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 21 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GACAAGGACG GGAGTGGGTA T 21

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 35 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GACTCGAGTC GACATCGATT TTTTTTTTTT TTTTT 35

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GACTCGAGTC GACATCG 17

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

AGGGGCTACG TAGTAAGGAC T 21

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ATTCTCAGGC TTAAGGTCCC T 21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 308 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Arg Asp Leu Lys Pro Glu Asn Phe Leu Phe Ser Ala Asp Asp Phe Met
 1               5                  10                  15
Val Lys Ser Lys Ala Thr Asp Phe Gly Leu Ser Val Phe Tyr Lys Pro
            20                  25                  30
Gly Gln Lys Phe Thr Asp Ile Val Gly Ser Pro Tyr Tyr Val Ala Pro
        35                  40                  45
Glu Val Leu Arg Lys Cys Tyr Gly Pro Gly Ser Asp Val Trp Ser Ala
    50                  55                  60
Gly Val Ile Leu Tyr Thr Leu Leu Cys Gly Ala Pro Pro Phe Met Ala
65                  70                  75                  80
Asp Ser Glu Pro Gly Val Ala Leu Gln Ile Leu His Gly Asp Leu Asp
                85                  90                  95
Phe Lys Ser Asp Pro Trp Pro Thr Ile Ser Glu Ser Ala Lys Asp Leu
            100                 105                 110
Ile Arg Lys Met Leu Glu Gln Asp Pro Lys Arg Arg Leu Thr Ala His
        115                 120                 125
Glu Val Leu Arg His Pro Trp Ile Val Asp Glu Asn Ile Ala Pro Asp
```

```
    130                 135                 140
Lys Pro Leu Gly Pro Ala Val Leu Ser Arg Leu Lys Gln Phe Ser Ala
145                 150                 155                 160
Met Asn Lys Ile Lys Lys Met Ala Leu Arg Val Ile Ala Glu Arg Leu
                165                 170                 175
Ser Glu Glu Glu Ile Val Gly Leu Lys Glu Met Phe Lys Met Ile Asp
            180                 185                 190
Thr Asp Asn Ser Gly Thr Val Thr Phe Phe His Leu Lys Asp Gly Leu
        195                 200                 205
Lys Arg Val Gly Ser Gln Leu Gly Glu Ser Glu Ile Lys Asp Leu Met
    210                 215                 220
Asp Ala Ala Asp Val Asp Asn Ser Gly Thr Ile Asp Tyr Gly Glu Phe
225                 230                 235                 240
Val Thr Ala Ala Met His Leu Asn Lys Ile Lys Arg Glu Asp His Leu
                245                 250                 255
Val Ser Ala Phe Ser Tyr His Asp Lys Asp Gly Ser Gly Tyr Ile Glu
            260                 265                 270
Val Asp Glu Ile Arg Gln Ala Leu Glu Glu Phe Gly Val Pro Asp Thr
        275                 280                 285
Ser Leu Glu Asp Met Ile Lys Glu Val Asp Thr Asp Asn Asp Gly Gln
    290                 295                 300
Ile Asp Tyr Gly
305
```

WHAT IS CLAIMED IS:

1. An isolated polynucleotide, said polynucleotide comprising:

a) the nucleotide sequence of SEQ ID NO:1;

b) an RNA analog of SEQ ID NO:1;

c) a polynucleotide comprising a nucleic acid sequence complementary to a) or b); or

d) a nucleic acid fragment of a), b) or c) that is at least 20 nucleotides in length and that hybridizes under stringent conditions to genomic DNA encoding the polypeptide of Figure 3.

2. The polynucleotide of claim 1, wherein said polynucleotide comprises nucleotides 1 to 170 of Figure 2.

3. The polynucleotide of claim 1, wherein said polynucleotide comprises nucleotides 160 to 560 of Figure 2.

4. The polynucleotide of claim 1, wherein said polynucleotide comprises nucleotides 550 to 920 of Figure 2.

5. A nucleic acid construct comprising the polynucleotide of claim 1.

6. The nucleic acid construct of claim 5, further comprising a regulatory element operably linked to said polynucleotide.

7. The nucleic acid construct of claim 6, wherein said regulatory element is an inducible regulatory element.

8. The nucleic acid construct of claim 7, wherein said regulatory element is induced in response to a plant pathogen.

9. A transgenic plant containing a nucleic acid construct comprising the polynucleotide of claim 1.

10. The plant of claim 9, wherein said construct further comprises a regulatory element operably linked to said polynucleotide.

11. The plant of claim 10, wherein said regulatory element is an inducible regulatory element.

12. The plant of claim 11, wherein said regulatory element is induced in response to a plant pathogen.

13. The plant of claim 11, wherein said regulatory element is induced in response to an elicitor.

14. The plant of claim 9, wherein said plant is a dicotyledonous plant.

15. The plant of claim 14, wherein said plant is a member of the *Solanaceae* family.

16. The plant of claim 15, wherein said plant is a *Nicotiana* plant.

17. The plant of claim 16, wherein said plant is *Nicotiana tabacum*.

18. A transgenic plant containing a polynucleotide expressing a polypeptide having from about

250 to about 550 amino acids, said polypeptide comprising an amino acid sequence substantially identical to the amino acid sequence of Figure 3.

19. The plant of claim 18, wherein said polypeptide comprises the amino acid sequence of Figure 3.

20. The plant of claim 18, wherein said plant is a dicotyledonous plant.

21. The plant of claim 20, wherein said plant is a member of the *Solanaceae* family.

22. A method of using a polynucleotide, said method comprising the step of hybridizing the polynucleotide of claim 1 to DNA or RNA from a plant.

23. The method of claim 22, further comprising the step of identifying a segment of said plant DNA or RNA that has about 70% or greater sequence identity to said polynucleotide.

24. The method of claim 23, further comprising the step of cloning at least a portion of said DNA or RNA segment.

25. The method of claim 24, wherein said cloned portion further comprises DNA flanking said segment having 70% or greater sequence identity.

26. A method of altering disease resistance in a plant, said method comprising the steps of:

(a) introducing the nucleic acid construct of claim 5 into a plant cell; and

(b) producing a plant containing said polynucleotide from said cell, wherein expression of said polynucleotide alters disease resistance in said plant.

27. The method of claim 26, wherein said nucleic acid construct further comprises an inducible regulatory element operably linked to said polynucleotide and said expression is regulated by said regulatory element.

28. The method of claim 27, wherein said expression is induced by said regulatory element upon exposure of said plant to an elicitor or plant pathogen.

29. An isolated polypeptide having from about 250 to about 550 amino acids, said polypeptide comprising an amino acid sequence substantially identical to Figure 3.

30. The polypeptide of claim 29, wherein said polypeptide comprises the amino acid sequence of Figure 3.

FIGURE 1

Primers

RecalV - 1 = GTTGACTCCCTACCCTCTT

↓

CALCIUM BINDING SITE

FokinB - 1 = GGTACTTAGGAAGTGTTACGGG

↓

KINASE DIAGNOSTIC SEQUENCE

FIGURE 2A

```
         10          20          30          40
          *           *           *           *
AGG GAC CTT AAG CCT GAG AAT TTC CTT TTC AGT GCC GAC GAC TTC
TCC CTG GAA TTC GGA CTC TTA AAG GAA AAG TCA CGG CTG CTG AAG

 50          60          70          80          90
  *           *           *           *           *
ATG GTA AAG AGT AAG GCC ACC GAC TTC GGG CTT AGT GTA TTC TAT
TAC CAT TTC TCA TTC CGG TGG CTG AAG CCC GAA TCA CAT AAG ATA

        100         110         120         130
          *           *           *           *
AAG CCT GGG CAA AAG TTC ACG GAC ATA GTA GGG AGT CCT TAC TAC
TTC GGA CCC GTT TTC AAG TGC CTG TAT CAT CCC TCA GGA ATG ATG

140         150         160         170         180
  *           *           *           *           *
GTA GCC CCT GAG GTA CTT AGG AAG TGT TAC GGG CCT GGG AGT GAC
CAT CGG GGA CTC CAT GAA TCC TTC ACA ATG CCC GGA CCC TCA CTG

        190         200         210         220
          *           *           *           *
GTA TGG AGT GCC GGG GTA ATA CTT TAC ACC CTT CTT TGT GGG GCC
CAT ACC TCA CGG CCC CAT TAT GAA ATG TGG GAA GAA ACA CCC CGG

230         240         250         260         270
  *           *           *           *           *
CCT CCT TTC ATG GCC GAC AGT GAG CCT GGG GTA GCC CTT CAA ATA
GGA GGA AAG TAC CGG CTG TCA CTC GGA CCC CAT CGG GAA GTT TAT

        280         290         300         310
          *           *           *           *
CTT CAT GGG GAC CTT GAC TTC AAG AGT GAC CCT TGG CCT ACC ATA
GAA GTA CCC CTG GAA CTG AAG TTC TCA CTG GGA ACC GGA TGG TAT

320         330         340         350         360
  *           *           *           *           *
AGT GAG AGT GCC AAG GAC CTT ATA AGG AAG ATG CTT GAG CAA GAC
TCA CTC TCA CGG TTC CTG GAA TAT TCC TTC TAC GAA CTC GTT CTG

        370         380         390         400
          *           *           *           *
CCT AAG AGG AGG CTT ACC GCC CAT GAG GTA CTT AGG CAT CCT TGG
GGA TTC TCC TCC GAA TGG CGG GTA CTC CAT GAA TCC GTA GGA ACC

410         420         430         440         450
  *           *           *           *           *
ATA GTA GAC GAG AAT ATA GCC CCT GAC AAG CCT CTT GGG CCT GCC
TAT CAT CTG CTC TTA TAT CGG GGA CTG TTC GGA GAA CCC GGA CGG

        460         470         480         490
          *           *           *           *
GTA CTT AGT AGG CTT AAG CAA TTC AGT GCC ATG AAT AAG ATA AAG
CAT GAA TCA TCC GAA TTC GTT AAG TCA CGG TAC TTA TTC TAT TTC

500         510         520         530         540
  *           *           *           *           *
```

FIGURE 2B

```
AAG ATG GCC CTT AGG GTA ATA GCC GAG AGG CTT AGT GAG GAG GAG
TTC TAC CGG GAA TCC CAT TAT CGG CTC TCC GAA TCA CTC CTC CTC

            550          560          570           580
             *            *            *             *
ATA GTA GGG CTT AAG GAG ATG TTC AAG ATG GAC ACC GAC AAT AGT
TAT CAT CCC GAA TTC CTC TAC AAG TTC TAC CTG TGG CTG TTA TCA

   590          600           610          620          630
    *            *             *            *            *
GGG ACC GTA ACC TTC TTC CAT CTT AAG CAA GGG CTT AAG AGG GTA
CCC TGG CAT TGG AAG AAG GTA GAA TTC GTT CCC GAA TTC TCC CAT

            640          650          660           670
             *            *            *             *
GGG AGT CAA CTT GGG GAG AGT GAG ATA AAG GAC CTT ATG GAC GCC
CCC TCA GTT GAA CCC CTC TCA CTC TAT TTC CTG GAA TAC CTG CGG

   680          690           700          710          720
    *            *             *            *            *
GCC GAC GTA GAC AAT AGT GGG ACC ATA GAC TAT GGG GAG TTC GTA
CGG CTG CAT CTG TTA TCA CCC TGG TAT CTG ATA CCC CTC AAG CAT

            730          740          750           760
             *            *            *             *
ACC GCC GCC ATG CAT CTT AAT AAG ATA AAG AGG GAG GAC CAT CTT
TGG CGG CGG TAC GTA GAA TTA TTC TAT TTC TCC CTC CTG GTA GAA

   770          780           790          800          810
    *            *             *            *            *
GTA AGT GCC TTC AGT TAT CAT GAC AAG GAC GGG AGT GGG TAT ATA
CAT TCA CGG AAG TCA ATA GTA CTG TTC CTG CCC TCA CCC ATA TAT

            820          830          840           850
             *            *            *             *
GAG GTA GAC GAG CTT AGG CAA GCC CTT GAG GAG TTC GGG GTA CCT
CTC CAT CTG CTC GAA TCC GTT CGG GAA CTC CTC AAG CCC CAT GGA

   860          870           880          890          900
    *            *             *            *            *
GAC ACC AGT CTT GAG GAC ATG ATA AAG GAG GTA GAC ACC GAC AAT
CTG TGG TCA GAA CTC CTG TAC TAT TTC CTC CAT CTG TGG CTG TTA

            910          920
             *            *
GAT GGG CAA ATA GAT TAT GGG
CTA CCC GTT TAT CTA ATA CCC
```

FIGURE 3

```
          10          20          30          40
           *           *           *           *
RDL KPE NLL FSA DDF MVK SKA TDF GLS VFY KPG QKF TDI VGS PYY

 50          60          70          80          90
  *           *           *           *           *
VAP EVL RKC YGP GSD VWS AGV ILY TLL CGA PPF MAD SEP GVA LQI

         100         110         120         130
           *           *           *           *
LHG DLD FKS DPW PTI SES AKD LIR KML EQD PKR RLT AHE VLR HPW

 140         150         160         170         180
  *           *           *           *           *
IVD ENI APD KPL GPA VLS RLK QFS AMN KIK KMA LRV IAE RLS EEE

         190         200         210         220
           *           *           *           *
IVG LKE MFK MDT DNS GTV TFF HLK QGL KRV GSQ LGE SEI KDL MDA

 230         240         250         260         270
  *           *           *           *           *
ADV DNS GTI DYG EFV TAA MHL NKI KRE DHL VSA FSY HDK DGS GYI

         280         290         300
           *           *           *
EVD ELR QAL EEF GVP DTS LED MIK EVD TDN DGQ IDY G
```

FIGURE 4

```
CDPKSOV    MAAKSSSSSTTTNVVTLKAAWVLPQRTQNIREVYEVGRKLGQGQFGTTFECTRRASGGKF

           ACKSIPKRKLLCKEDYEDVWREIQIMHHLSEHANVVRIEGTYEDSTAVHLVMELCEGGEL

           FDRIVQKGHYSERQAARLIKTIVEVVEACHSLGVMHRDLKPENFLFDTIDEDAKLKATDF
                                                .......  ..SAD.FMV.S.....

           GLSVFYKPGESFCDVVGSPYYVAPEVLRKLYGPESDVWSAGVILYILLSGVPPFWAESEP
           .........QK.T.I.........C...G...........T..C.A...M.D...

           GIFRQILLGKLDFHSEPWPSISDSAKDLIRKMLDQNPKTRLTAHEVLRHPWIVDDNIAPD
           .VAL...H.D...K.D...T..E..........E.D..R...............E.....

           KPLDSAVLSRLKQFSAMNKLKKMALRVIAERLSEEEIGGLKELFKMIDTDNSGTITFDEL
           ...GP..............I.................V....M...-.......V..FH.

           KDGLKRVGSELMESEIKDLMDAADIDKSGTIDYGEFIAATVHLNKLEREENLVSAFSYFD
           .........Q.G...........V.N.......  .VT.AM....IK..DH.......H.

           KDGSGYITLDEIQQACKDFGLDDIHIDDMIKEIDQDNDGQIDYGEFAAMMRKGNGGIGRR
           .......EV..R..LEE..VP.TSLE.....VT.........

           TMRKTLNLRDALGLVDNGSNQVIEGYFK
```

→ PRIMERS

PROTEIN KINASE SEQUENCES

CALCIUM-BINDING SITES

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/14109

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C12N 5/04, 5/10, 5/16, 5/22, 15/82; A01H 5/00
US CL : 435/34; 800/278,530/300
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/34; 800/278,530/300

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS. MEDLINE, BIOSIS, AGRICOLA

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X - Y	HARPER et al. A Calcium-Dependent Protein Kinase with a Regulatory Domain Similar to Calmodulin. Science. 17 May 1991, Vol. 252, page 951-954, see entire document.	1 — 2-21, 26-28
Y	VALVEKENS et al. Agrobacterium Tumefaciens-Mediated Transformation of Arabidopsis Thaliana Root Explants by Using Kanamycin Selection. Proc. Natl. Acad. Sci. August 1988, Vol. 85, pages 5536-5540, see entire document.	2-21,26-28

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
02 OCTOBER 1998	23 OCT 1998
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231	Authorized officer Dorothea Lawrence for OUSAMA M-FAIZ ZAGHMOUT
Facsimile No. (703) 305-3230	Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)★

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/14109

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-21,26-28

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.
Group I. Claims 1-21 and 26-28 are drawn to nucleic acid molecule encoding CDPK from tobacco, vectors containing it in sense orientation, methods for their use to transform plants, and the resultant plants.

Group II. Claims 22-25 are drawn to the use of polynucleotides to isolate DNA fragment comprising less than full length gene.

Group III. Claims 29-30 are drawn to CDPK protein.

The inventions listed as groups I-III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: Since fragment of nucleotide sequence of the protein is known in the art as evidenced by the Harper et al reference (Science. 1991. Vol. 252:951-954), it does not constitute a special technical feature as defined by PCT Rule 13.2. Groups I-III are directed to isolation and use of nucleic acid from plant cells and their expression in transgenic plants in addition to the CDPK protein of group III. However, since claim 1 lacks novelty, unity of invention is lacking, because fragment of nucleotide sequence of the protein was reported previously by the Harper et al reference (Science. 1991. Vol. 252:951-954). The cited evidence proves that the technical feature of group 1, fragment of nucleotide sequence of the protein, does not make a contribution over the prior art. The claims are not so linked by a special technical feature within the meaning of the PCT Rule 13.2 so as to form a single inventive concept, accordingly, the unity of invention is lacking among all groups.